JOURNAL OF MULTIVARIATE ANALYSIS 38, 187–203 (1991)

An Empirical Process Central Limit Theorem for Dependent Non-identically Distributed Random Variables

DONALD W. K. ANDREWS*

Yale University

Communicated by C. R. Rao

This paper establishes a central limit theorem (CLT) for empirical processes indexed by smooth functions. The underlying random variables may be temporally dependent and non-identically distributed. In particular, the CLT holds for near epoch dependent (i.e., functions of mixing processes) triangular arrays, which include strong mixing arrays, among others. The results apply to classes of functions that have series expansions. The proof of the CLT is particularly simple; no chaining argument is required. The results can be used to establish the asymptotic normality of semiparametric estimators in time series contexts. An example is provided.

1. INTRODUCTION

This paper establishes a central limit theorem (CLT) for empirical processes indexed by classes of smooth functions. The contribution of the paper is twofold. First, the results improve existing results by allowing the underlying random variables (rv's) to possess a more general form of temporal dependence and non-identical distributions than is available elsewhere. Second, the proof of the results is very simple.

The results of this paper apply to near epoch dependent (NED) (i.e., functions of mixing processes) triangular arrays of rv's. In contrast, existing results in the empirical process literature only consider independent

Received June 6, 1989; revised July 31, 1990.

AMS classification numbers: 60F05, 60F17.

Key words and phrases: central limit theorem, empirical process, Fourier series, functional central limit theorem, near epoch dependence, semiparametric estimator, series expansion, Sobolev norm, stochastic equicontinuity, strong mixing.

* I thank David Pollard, Yoon-Jae Whang, and a referee for helpful comments. I gratefully acknowledge research support from the Alfred P. Sloan Foundation and the National Science Foundation through a Research Fellowship and Grants SES-8618617 and SES-8821021, respectively.

0047-259X/91 $3.00

rv's or strong mixing rv's (for the latter, see Andrews and Pollard [4]), while results in the Banach space-valued CLT literature only consider independent or weakly stationary strong mixing sequences of rv's (e.g., see Dehling and Philipp [9] and Dehling [8]). NED triangular arrays include strong mixing triangular arrays, as well as a rich class of non-strong mixing triangular arrays.

The approach used here is to consider an index class of functions that have series expansions with respect to the same set of basis functions. A prime example is a class of functions that have Fourier series expansions. Conditions are placed on the coefficients of the series expansions that are sufficient to obtain stochastic equicontinuity of the empirical process and a CLT for it. In particular applications, these conditions can be verified by using known results for series expansions. The case of differentiable functions on an open bounded subset of R^k is examined in detail. In this case, if the rv's are strong mixing, the required number of finite derivatives of the functions is the same as for Ossiander's [19] results for i.i.d. rv's obtained using a bracketing condition.

Note that the series expansion approach used here is different from approaches currently used in the empirical process literature, but is similar to an approach used in the Banach space-valued CLT literature (see the references above).

An appealing feature of the series expansion approach is the simplicity with which one obtains the stochastic equicontinuity and CLT results. Elementary manipulations and inequalities suffice (see the proof of Theorem 1). No chaining argument is required. In consequence, the series expansion approach is conducive to obtaining results under quite weak assumptions regarding temporal dependence and non-identical distributions (as noted aove). The drawback of this approach is that the variety of different classes of functions to which it applies is more restricted than with a bracketing approach (e.g., see Ossiander [19] or Pollard [20]) or a Vapnik–Cervonenkis approach (e.g., see Pollard [21]).

The empirical process results of this paper have numerous applications in econometrics and statistics. For example, they are tailor-made for use with the results of Andrews [2, 3], which establish the $\sqrt{n}$-consistency and asymptotic normality of semiparametric and parametric estimators and the asymptotic chi-squared distributions of Wald, Lagrange multiplier, and likelihood ratio-like test statistics that correspond to such estimators.

The remainder of this paper is organized as follows: Section 2 provides the main CLT results. Section 3 applies the CLT results to the case of indexing by a class of differentiable functions. Section 4 illustrates the use of the results in a semiparametric estimation problem.

2. An Empirical Process CLT

2.1. *Notation and Definitions*

Let $\{X_{ni}: 1 \leqslant i \leqslant n, n \geqslant 1\}$ be a triangular array of $\mathscr{X}$-valued random vectors (rv's) defined on a probability space $(\Omega, \mathscr{B}, \mathbf{P})$, where $\mathscr{X} \subset R^k$. Let $m(\cdot, \cdot)$ be a real function defined on $\mathscr{X} \times \mathscr{T}$, where $\mathscr{T}$ is an index set that is a metric space with metric ρ (defined below). Assume $m(x, \tau)$ is Borel measurable in x for each fixed $\tau \in \mathscr{T}$. Define the empirical process $v_n(\cdot)$ by

$$v_n(\tau) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (m(X_{ni}, \tau) - Em(X_{ni}, \tau)). \tag{2.1}$$

We define weak convergence (denoted by w) as in Dudley [10] and Hoffman-Jorgensen [14] (with a slight, inessential but convenient, modification introduced by Pollard [21]). Let $L^\infty(\mathscr{T})$ denote the space of bounded real functions on $\mathscr{T}$. Endow $L^\infty(\mathscr{T})$ with the uniform metric d.

Definition. If $\{v_n(\cdot): n \geqslant 1\}$ are (not necessarily Borel measurable) maps from Ω into the metric space $(L^\infty(\mathscr{T}), d)$ and if $v(\cdot)$ is an $L^\infty(\mathscr{T})$-valued Borel measurable rv (not necessarily defined on $(\Omega, \mathscr{B}, \mathbf{P})$), then

$$v_n(\cdot)\, w\, v(\cdot) \qquad \text{if} \quad E^* f(v_n(\cdot)) \to Ef(v(\cdot)) \qquad \text{as} \quad n \to \infty \tag{2.2}$$

for all bounded uniformly continuous real functions f on $L^\infty(\mathscr{T})$, where E^* denotes outer expectation.

Definition. $\{v_n(\cdot): n \geqslant 1\}$ is stochastically equicontinuous if $\forall \varepsilon > 0$ and $\eta > 0$, $\exists \delta > 0$ such that

$$\overline{\lim_{n \to \infty}}\, \mathbf{P}^*\Big(\sup_{\rho(\tau, \gamma) < \delta} |v_n(\tau) - v_n(\gamma)| > \eta \Big) < \varepsilon, \tag{2.3}$$

where $\mathbf{P}^*$ denotes $\mathbf{P}$-outer probability.

The process $v_n(\cdot)$ converges weakly to a Gaussian process if it is stochastically equicontinuous, its finite dimensional distributions are asymptotically normal, and $(\mathscr{T}, \rho)$ is totally bounded (see Pollard [21, Theorem 10.2]). In view of this result, we organize this section as follows. Subsection 2.2 gives general conditions under which $\{v_n(\cdot): n \geqslant 1\}$ is stochastically equicontinuous. Subsection 2.3 uses these conditions to obtain conditions for stochastic equicontinuity when the rv's $\{X_{ni}\}$ are NED. Subsection 2.4 provides conditions under which the finite dimensional distributions of $\{v_n(\cdot): n \geqslant 1\}$ are asymptotically normal when the rv's $\{X_{ni}\}$ are NED. Subsection 2.5 states the main CLT result.

2.2. *Stochastic Equicontinuity*

The following assumption implies that $\{v_n(\cdot): n \geqslant 1\}$ is stochastically equicontinuous.

Assumption A. (i) (Series expansion). For some sequence $\{h_j(\cdot): j \geqslant 1\}$ of real or complex Borel measurable functions on $\mathscr{X}$, $m(\cdot, \tau)$ has a pointwise convergent series expansion for each $\tau \in \mathscr{T}$: $m(x, \tau) = \sum_{j=1}^{\infty} c_j(\tau)\, h_j(x)$ $\forall x \in \mathscr{X}$, where for each $\tau \in \mathscr{T}$, $\{c_j(\tau): j \geqslant 1\}$ is a sequence of (real or complex) constants.

(ii) (Smoothness). $\sum_{j=1}^{\infty} |c_j(\tau)|\, E\,|h_j(X_{ni})| < \infty$ $\forall i \leqslant n, n \geqslant 1, \tau \in \mathscr{T}$.

(iii) (Smoothness/weak dependence trade-off). $\sup_{\tau \in \mathscr{T}} \sum_{j=J}^{\infty} |c_j(\tau)|^2/a_j \to 0$ as $J \to \infty$ for some summable sequence of positive real constants $\{a_j\}$ for which $\sum_{j=1}^{\infty} a_j \gamma_j < \infty$, where $\gamma_j = \sum_{s=-\infty}^{\infty} \gamma_j(s)$ and $\gamma_j(s) = \sup_{i \leqslant n-|s|, n \geqslant 1} |\mathrm{Cov}(h_j(X_{ni}), h_j(X_{ni+|s|}))|$.

The series functions $\{h_j(\cdot)\}$ of Assumption A(i) need not be orthogonal or orthonormal, nor do the functions $\{m(\cdot, \tau)\}$ need to have unique series expansions in terms of $\{h_j(\cdot)\}$. All that is required is that one can identify a single sequence of coefficients $\{c_j(\tau)\}$ (out of many perhaps) with each function $m(\cdot, \tau)$.

Since each $\tau \in \mathscr{T}$ is identified with a particular sequence $\{c_j(\tau): j \geqslant 1\}$ under Assumption A, the metric ρ on $\mathscr{T}$ can be taken to be

$$\rho(\tau, \gamma) = \left(\sum_{j=1}^{\infty} |c_j(\tau) - c_j(\gamma)|^2 \right)^{1/2}, \qquad \forall \tau, \gamma \in \mathscr{T}. \tag{2.4}$$

THEOREM 1. *For ρ as defined in* (2.4), *Assumption A implies that $\{v_n(\cdot)\}$ is stochastically equicontinuous and $(\mathscr{T}, \rho)$ is totally bounded.*

Proof of Theorem 1. First we show stochastic equicontinuity. By Assumptions A(i) and A(ii), $m(X_{ni}, \tau) - Em(X_{ni}, \tau) = \sum_{j=1}^{\infty} c_j(\tau)(h_j(X_{ni}) - Eh_j(X_{ni}))$. Thus, we have

$$\begin{aligned}
&\overline{\lim_{n\to\infty}}\ \mathbf{P}^*\Big(\sup_{\rho(\tau,\gamma)<\delta} |v_n(\tau) - v_n(\gamma)| > \eta \Big) \\
&\quad \leqslant \overline{\lim_{n\to\infty}}\ \eta^{-2} E^* \Bigg(\sup_{\rho(\tau,\gamma)<\delta} \Bigg| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{\infty} (c_j(\tau) - c_j(\gamma)) \\
&\qquad \times (h_j(X_{ni}) - Eh_j(X_{ni})) \Bigg| \Bigg)^2 \\
&\quad \leqslant \eta^{-2} \sup_{\rho(\tau,\gamma)<\delta} \sum_{j=1}^{\infty} \frac{|c_j(\tau) - c_j(\gamma)|^2}{a_j}\ \overline{\lim_{n\to\infty}}\ E^* \sum_{j=1}^{\infty} a_j \\
&\qquad \times \Bigg| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (h_j(X_{ni}) - Eh_j(X_{ni})) \Bigg|^2 \\
&\quad \leqslant \eta^{-2} \sup_{\rho(\tau,\gamma)<\delta} \sum_{j=1}^{\infty} |c_j(\tau) - c_j(\gamma)|^2/a_j \cdot \sum_{j=1}^{\infty} a_j \gamma_j,
\end{aligned} \tag{2.5}$$

where the second inequality uses the Cauchy–Schwarz inequality and the third uses standard manipulations of the variance of a sum of correlated rv's.

Stochastic equicontinuity follows from (2.5) and

$$\lim_{\delta \to 0} \sup_{\rho(\tau,\gamma)<\delta} \sum_{j=1}^{\infty} |c_j(\tau) - c_j(\gamma)|^2/a_j = 0. \tag{2.6}$$

To obtain (2.6), suppose $\varepsilon > 0$ is given. By Assumption A(iii), one can choose J sufficiently large that $\sup_{\tau,\gamma \in \mathscr{T}} \sum_{j=J+1}^{\infty} |c_j(\tau) - c_j(\gamma)|^2/a_j < \varepsilon/2$. Take, $\delta = [\varepsilon/(2\sum_{j=1}^{J} 1/a_j)]^{1/2}$. By definition of $\rho(\cdot,\cdot)$, this gives $\sup_{\rho(\tau,\gamma)<\delta} \sum_{j=1}^{J} |c_j(\tau) - c_j(\gamma)|^2/a_j \leqslant \varepsilon/2$, which establishes (2.6).

Next, we establish totally boundedness of $(\mathscr{T}, \rho)$. By Assumption A(iii), $\forall \varepsilon > 0$, $\exists J < \infty$ such that $\sup_{\tau \in \mathscr{T}} \sum_{j=J+1}^{\infty} |c_j(\tau)|^2 < \varepsilon/2$. In addition, $\sup_{\tau \in \mathscr{T}} |c_j(\tau)|^2 < K$ for some $K < \infty$ $\forall j = 1, ..., J$. The result follows by showing that the set of functions $\{\sum_{j=1}^{J} c_j h_j(\cdot) : |c_j|^2 \leqslant K \, \forall j \leqslant J\}$ can be covered by a finite number of $\varepsilon/2$-balls. ∎

2.3. *Stochastic Equicontinuity with NED Random Vectors*

This subsection replaces the weak dependence condition of Assumption A(iii) by a more standard condition of weak temporal dependence, viz., the NED condition. The NED condition was introduced by Ibragimov [15] and results utilizing it were developed by Billingsley [6, p. 182] and McLeish [17, 18]. The NED condition is quite general. It allows for non-identical distributions and covers (i) square integrable strong mixing rv's, (ii) square integrable general linear processes (with non-identically distributed strong mixing innovations if desired) including autoregressive and autoregressive-moving average processes (which are not necessarily strong mixing, e.g., see Andrews [1]), and (iii) various nonlinear autoregressions and dynamic simultaneous equations (see Bierens [5, Chap. 5], Gallant [12, pp. 481, 502, 539], and Gallant and White [13, p. 29]).

First we define strong mixing double arrays and NED triangular arrays of rv's. Let $\{V_{ni} : i = 0, \pm 1, ...; n \geqslant 1\}$ be a double array of rv's on $(\Omega, \mathscr{B}, \mathbf{P})$. ($V_{ni}$ may be $\mathscr{V}_{ni}$-valued for any measurable space $\mathscr{V}_{ni}$, but usually $\mathscr{V}_{ni} = \mathscr{V}$ and $\mathscr{V} \subset R^v$ or $\mathscr{V} \subset \mathscr{C}^v$ for some $v \geqslant 1$, where $\mathscr{C}$ denotes the complex plane.) Let $\mathscr{F}_{n,i}^{j}$ $(\subset \mathscr{B})$ denote the σ-field generated by $(V_{ni}, ..., V_{nj})$ for $-\infty \leqslant i \leqslant j \leqslant \infty$. Let $E_{n,i}^{j}(\cdot)$ denote the conditional expectation operator $E(\cdot \mid \mathscr{F}_{n,i}^{j})$ given the σ-field $\mathscr{F}_{n,i}^{j}$.

DEFINITION. The double array $\{V_{ni}\}$ of rv's is strong mixing if $\alpha(s) \downarrow 0$ as $s \to \infty$, where

$$\alpha(s) = \sup_{i=0,\pm 1,...;n \geqslant 1} \; \sup_{A \in \mathscr{F}_{n,-\infty}^{i},\, B \in \mathscr{F}_{n,i+s}^{\infty}} |\mathbf{P}(A \cap B) - \mathbf{P}(A)\,\mathbf{P}(B)| \quad \text{for} \quad s \geqslant 1. \tag{2.7}$$

$\{V_{ni}\}$ is strong mixing of size $-\beta$ if $\alpha(s) = O(s^{-\beta-\varepsilon})$ for some $\varepsilon > 0$.

Let $\{Z_{ni}: i \leqslant n, n \geqslant 1\}$ be a triangular array of R^k- or $\mathscr{C}^k$-valued rv's on $(\Omega, \mathscr{B}, \mathbf{P})$. Let $\|\cdot\|$ denote the Euclidean norm.

DEFINITION. The triangular array $\{Z_{ni}\}$ is near epoch dependent (NED) on $\{V_{ni}\}$ if $E\|Z_{ni}\|^2 < \infty\ \forall i \leqslant n, \forall n \geqslant 1$, and $\eta(s) \downarrow 0$ as $s \to \infty$, where

$$\eta(s) = \sup_{s < i \leqslant n-s, n \geqslant 1} (E\|Z_{ni} - E_{n,i-s}^{i+s} Z_{ni}\|^2)^{1/2} \quad \text{for} \quad s = 0, 1, ... \tag{2.8}$$

$\{Z_{ni}\}$ is NED of size $-\beta$ on $\{V_{ni}\}$ if $E\|Z_{ni}\|^2 < \infty$, $\forall i \leqslant n, n \geqslant 1$ and $\eta(s) = O(s^{-\beta-\varepsilon})$ for some $\varepsilon > 0$.

(It is easy to see that the R^k- or $\mathscr{C}^k$-valued array $\{Z_{ni}\}$ is NED of size $-\beta$ on $\{V_{ni}\}$ if and only if the real$-$ or complex-valued array $\{Z_{nil}\}$ is NED of size $-\beta$ on $\{V_{ni}\}$ for $l = 1, ..., k$, where $Z_{ni} = (Z_{ni1}, ..., Z_{nik})'$.)

DEFINITION. Let $\mathrm{Lip}(\alpha, C, \mathscr{X})$ denote the class of real or complex functions g on $\mathscr{X} \subset R^k$ that satisfy the Lipschitz condition $|g(x) - g(y)| \leqslant C\|x - y\|^{\alpha}\ \forall x, y \in \mathscr{X}$, where $0 < C < \infty$ and $0 < \alpha \leqslant 1$.

The following assumption implies Assumption A and, hence, is sufficient for stochastic equicontinuity of $\{v_n(\cdot): n \geqslant 1\}$.

Assumption A1. For some $r > 2$,

(i) (Series expansion). Assumption A(i) holds,

(ii) (Smoothness of series functions). $h_j(\cdot) \in \mathrm{Lip}(1, B_j, \mathscr{X})$ for some $B_j < \infty\ \forall j \geqslant 1$ and $\sup_{j \geqslant 1} \sup_{i \leqslant n, n \geqslant 1} E|h_j(X_{ni})|^r < \infty$,

(iii) (Smoothness of $m(\cdot, \tau)$). $\sup_{\tau \in \mathscr{T}} \sum_{j=J}^{\infty} |c_j(\tau)|^2/a_j \to 0$ as $J \to \infty$ for some summable sequence of positive constants $\{a_j\}$ for which $\sum_{j=1}^{\infty} a_j B_j < \infty$, and

(iv) (Weak dependence). $\{X_{ni}: i \leqslant n, n \geqslant 1\}$ is a NED triangular array of size -1 on $\{V_{ni}\}$, where $\{V_{ni}\}$ is some strong mixing double array of size $-2r/(r-2)$.

THEOREM 2. *Assumption* A1 *implies Assumption* A.

To prove Theorem 2 we use the following two lemmas.

LEMMA 1. *If $g \in \mathrm{Lip}(1, C, \mathscr{X})$ and $\{Z_{ni}: i \leqslant n, n \geqslant 1\}$ is a NED triangular array of $\mathscr{X}$-valued rv's of size $-\xi$ on a strong mixing double array $\{V_{ni}\}$ of size $-\beta$ for $\xi, \beta > 0$, then $\{g(Z_{ni}): i \leqslant n, n \geqslant 1\}$ is a NED triangular array of real- or complex-valued rv's of size $-\xi$ on the same strong mixing double array $\{V_{ni}\}$. In addition, the NED numbers $\{\eta_g(s)\}$ of $\{g(Z_{ni})\}$ satisfy $\eta_g(s) \leqslant C\eta(s)$ for all $s \geqslant 0$, where $\{\eta(s)\}$ are the NED numbers of $\{Z_{ni}\}$.*

This lemma is similar to Theorem 4.2 of Gallant and White [13, p. 48]. It differs from the latter in that Gallant and White consider more general functions than those in Lip$(1, C, \mathscr{X})$, but their result does not obtain the same size of the NED numbers of $\{g(Z_{ni})\}$ as those of $\{Z_{ni}\}$ and their result imposes stronger moment conditions.

Proof of Lemma 1. $E\,|g(Z_{ni})|^2 < \infty$ by the Lipschitz condition and $E\,\|Z_{ni}\|^2 < \infty$.

Next, note that the conditional expectation $E_{n,i-s}^{i+s}\, g(Z_{ni})$ minimizes $E_{n,i-s}^{i+s}\,|g(Z_{ni}) - Y|^2$ over all $\mathscr{F}_{n,i-s}^{i+s}$-measurable rv's Y. With this result and the Lipschitz condition on g, we obtain

$$\begin{aligned}\eta_g(s) &= \sup_{s<i\leqslant n-s,\, n\geqslant 1} (EE_{n,i-s}^{i+s}\,|g(Z_{ni}) - E_{n,i-s}^{i+s}\, g(Z_{ni})|^2)^{1/2}\\ &\leqslant \sup_{s<i\leqslant n-s,\, n\geqslant 1} (EE_{n,i-s}^{i+s}\,|g(Z_{ni}) - g(E_{n,i-s}^{i+s} Z_{ni})|^2)^{1/2}\\ &\leqslant C \sup_{s<i\leqslant n-s,\, n\geqslant 1} (E\|Z_{ni} - E_{n,i-s}^{i+s} Z_{ni}\|^2)^{1/2} = C\eta(s). \quad \blacksquare \end{aligned} \tag{2.9}$$

Let $\|\cdot\|_p$ denote the $L^p(\mathbf{P})$ semi-norm for $p \geqslant 1$.

LEMMA 2. *Let $\{Z_{ni}\}$ be a NED triangular array on $\{V_{ni}\}$ (with NED numbers $\{\eta(s): s\geqslant 0\}$), where $\{V_{ni}\}$ is a strong mixing double array (with mixing numbers $\{\alpha(s): s\geqslant 1\}$). Then, for all $r>2$, $|\mathrm{Cov}(Z_{ni}, Z_{ni-s})| \leqslant \|Z_{ni-s}\|_2\, \eta(a) + 6\, \|Z_{ni}\|_2\, \|Z_{ni-s}\|_r\, \alpha(b)^{(r-2)/(2r)}$ for all $0 \leqslant s < i$ and $n \geqslant 1$, where a and b are positive integers for which $a + b \leqslant s$.*

Proof of Lemma 2. We have

$$\begin{aligned}|\mathrm{Cov}(Z_{ni}, Z_{ni-s})| &= |E(Z_{ni} - E_{n,i-a}^{i+a} Z_{ni})(Z_{ni-s} - EZ_{ni-s})\\ &\quad + E(E_{n,i-a}^{i+a} Z_{ni})\, E_{n,i-a}^{i+a}(Z_{ni-s} - EZ_{ni-s})|\\ &\leqslant (E\,|Z_{ni} - E_{n,i-a}^{i+a} Z_{ni}|^2)^{1/2}\, \mathrm{Var}^{1/2}(Z_{ni-s})\\ &\quad + (E\,|E_{n,i-a}^{i+a} Z_{ni}|^2)^{1/2}\, (E\,|E_{n,i-a}^{i+a} Z_{ni-s} - EZ_{ni-s}|^2)^{1/2}\\ &\leqslant \eta(a)\, \|Z_{ni-s}\|_2 + \|Z_{ni}\|_2\, 6\alpha(b)^{(r-2)/(2r)} \|Z_{ni-s}\|_r, \end{aligned} \tag{2.10}$$

using the Cauchy–Schwarz inequality, the conditional Jensen's inequality, and a strong mixing inequality of McLeish [17, Lemma 2.1]. $\blacksquare$

Proof of Theorem 2. Assumptions A1(i) and A(i) are equivalent. By Assumption A1(ii), Assumption A(ii) holds if $\sum_{j=1}^{\infty} |c_j(\tau)| < \infty$ $\forall \tau \in \mathscr{T}$. This holds under Assumption A1(iii) because $\sum_{j=1}^{\infty} |c_j(\tau)| \leqslant (\sum_{j=1}^{\infty} |c_j(\tau)|^2/a_j)^{1/2}\, (\sum_{j=1}^{\infty} a_j)^{1/2} < \infty$.

Next we show that Assumption A1 implies A(iii). By Assumptions A1(ii) and (iv) and Lemma 1, $\{h_j(X_{ni})\}$ is a NED triangular array on $\{V_{ni}\}$ with NED numbers $\{\eta_{h_j}(s)\}$ that satisfy $\eta_{h_j}(s) \leqslant B_j \eta(s)\ \forall s \geqslant 0$, where $\{\eta(s)\}$ are the NED numbers of $\{X_{ni}\}$. Thus, Assumption A1(ii) and Lemma 2 with $Z_{ni} = h_j(X_{ni})$ and $a = b = [s/2]$ give

$$\gamma_j(s) \leqslant \sup_{\substack{i \leqslant n, n \geqslant 1 \\ v \geqslant 1}} \|h_v(X_{ni})\|_2 \,(B_j \eta([s/2])$$

$$+ 6 \sup_{\substack{i \leqslant n, n \geqslant 1 \\ v \geqslant 1}} \|h_v(X_{ni})\|_r \,\alpha([s/2])^{(r-2)/(2r)}),$$

$$\gamma_j = \sum_{s=-\infty}^{\infty} \gamma_j(s) \leqslant B_j D_1 \sum_{s=-\infty}^{\infty} \eta([s/2])$$

$$+ D_2 \sum_{s=-\infty}^{\infty} \alpha([s/2])^{(r-2)/(2r)} < \infty \qquad (2.11)$$

$\forall j \geqslant 1$ for some finite constants D_1 and D_2, using the fact that $\{\eta(s)\}$ and $\{\alpha(s)\}$ are of size -1 and $-2r/(r-2)$, respectively. Thus, $\sum_{j=1}^{\infty} a_j \gamma_j < \infty$ is implied by $\sum_{j=1}^{\infty} a_j B_j < \infty$ and $\sum_{j=1}^{\infty} a_j < \infty$. This result and Assumption A1(iii) imply Assumption A(iii). ∎

2.4. *Fidi Convergence*

We now prove a set of primitive conditions under which the finite dimensional distributions of $v_n(\cdot)$ are asymptotically normal. To this end, we state a CLT of Wooldridge [25] for triangular arrays of rv's that are NED.

PROPOSITION 1 (Wooldridge [25, Chap. 2, Theorem 3.13]). *Let $\{Z_{ni}: i \leqslant n, n \geqslant 1\}$ be a triangular array of real-valued rv's that satisfies*

(i) $\lim_{n\to\infty} \mathrm{Var}((1/\sqrt{n}) \sum_{i=1}^{n} Z_{ni}) = \sigma^2$,

(ii) $\sup_{i \leqslant n, n \geqslant 1} E\,|Z_{ni}|^r < \infty$ *for some* $r > 2$, *and*

(iii) $\{Z_{ni}\}$ *is NED of size* -1 *on* $\{V_{ni}\}$, *where* $\{V_{ni}\}$ *is a strong mixing double array of rv's of size* $-2r/(r-2)$.

Then, $(1/\sqrt{n}) \sum_{i=1}^{n} (Z_{ni} - EZ_{ni})\, dN(0, \sigma^2)$ *as* $n \to \infty$.

Comment. Wooldridge's proof uses the approach of Withers [24].

The following assumption implies fidi convergence for $v_n(\cdot)$:

Assumption B (Fidi convergence). For some $r > 2$,

(i) $S(\tau, \gamma) = \lim_{n\to\infty} \mathrm{Cov}(v_n(\tau), v_n(\gamma))$ exists $\forall \tau, \gamma \in \mathscr{T}$,

(ii) $\sup_{i \leqslant n, n \geqslant 1} E\,\|X_{ni}\|^r < \infty$,

(iii) $\{X_{ni}: i \leqslant n, n \geqslant 1\}$ is a NED triangular array of size -1 on $\{V_{ni}\}$, where $\{V_{ni}: i = 0, \pm 1, ...; n \geqslant 1\}$ is some strong mixing double array of size $-2r/(r-2)$, and

(iv) $m(\cdot, \tau) \in \mathrm{Lip}(1, C, \mathscr{X})\ \forall \tau \in \mathscr{T}$ for some $C < \infty$.

THEOREM 3. *Under Assumption* B, *for each finite subset* $(\tau_1, ..., \tau_v)$ of $\mathscr{T}$, $(v_n(\tau_1), ..., v_n(\tau_v))'$ *converges in distribution to a* $N(\mathbf{0}, S_v)$ *rv, where* S_v *is a* $v \times v$ *covariance matrix with* (s, t)*th element* $S(\tau_s, \tau_t)$.

Proof of Theorem 3. Let $\boldsymbol{\tau} = (\tau_1, ..., \tau_v)' \in \mathscr{T}^v$ and $m(\cdot, \boldsymbol{\tau}) = (m(\cdot, \tau_1), ..., m(\cdot, \tau_v))'$. It suffices to show that conditions (i), (ii), and (iii) of Proposition 1 hold with $Z_{ni} = \lambda' m(X_{ni}, \boldsymbol{\tau})$ and $\sigma^2 = \lambda' S_v \lambda$, $\forall \boldsymbol{\tau} \in \mathscr{T}^v$, $\forall \lambda \in R^v$ with $\|\lambda\| = 1$, and $\forall v \geqslant 1$.

Assumption B(i) implies condition (i) of Proposition 1. In addition, Assumptions B(ii) and (iv) imply that $\sup_{i \leqslant n, n \geqslant 1} E|\lambda' m(X_{ni}, \boldsymbol{\tau})|^r < \infty$. That is, condition (ii) of Proposition 1 holds. Assumptions B(iii) and (iv) and Lemma 1 with $g(\cdot) = \lambda' m(\cdot, \boldsymbol{\tau})$ imply that $\{\lambda' m(X_{ni}, \boldsymbol{\tau})\}$ is NED of size -1 on $\{V_{ni}\}$. That is, condition (iii) of Proposition 1 holds. ∎

2.5. *An Empirical Process CLT*

Theorems 1–3 above combine to yield a CLT for $\{v_n(\cdot): n \geqslant 1\}$. Define

$$\mathscr{U}_\rho(\mathscr{T}) = \{y \in L^\infty(\mathscr{T}): y \text{ is uniformly continuous with respect to } \rho \text{ on } \mathscr{T}\}. \tag{2.12}$$

COROLLARY 1. *For* $v_n(\cdot)$ *and* ρ *as defined in* (2.1) *and* (2.4), *Assumption* A *or* A1 *plus Assumption B imply* $v_n(\cdot)\, w v(\cdot)$, *where* $v(\cdot)$ *is a mean zero Gaussian process with covariance function* $S(\cdot, \cdot)$ *whose sample paths lie in* $\mathscr{U}_\rho(\mathscr{T})$ *with probability one.*

Proof of Corollary 1. Theorems 1–3 and Pollard [21, Theorem 10.2] give the result. ∎

3. DIFFERENTIABLE FUNCTIONS OF R^k-VALUED RANDOM VARIABLES

In this section we apply the results of Section 2 to a class of differentiable functions $\{m(\cdot, \tau): \tau \in \mathscr{T}\}$ that are defined on some open bounded subset $\mathscr{X}$ of R^k whose boundary is *minimally smooth.* (See Stein [23, pp. 181, 189] or Edmunds and Moscatelli [11, p. 8] for the definition of minimally smooth.) Examples of sets in R^k with minimally smooth boundaries include open sets that are convex or whose boundaries are C^1-embedded in R^k. Finite unions of disjoint sets of the aforementioned type also have minimally smooth boundaries.

The functions $\{m(\cdot, \tau)\}$ are taken to be uniformly smooth in the sense of having a uniformly bounded Sobolev norm of some order. By definition, the Sobolev norm of order (q, p) of a real function f defined on a subset $\mathscr{Y}$ of R^k is

$$\|f\|_{q,p,\mathscr{Y}} = \left(\sum_{|\alpha| \leqslant q} \int_{\mathscr{Y}} |D^\alpha f(x)|^p \, dx \right)^{1/p}, \tag{3.1}$$

where q is a non-negative integer, $1 \leqslant p < \infty$, $\alpha = (\alpha_1, ..., \alpha_k)' \in R^k$ has non-negative integer-valued elements, $|\alpha| = \sum_{l=1}^k \alpha_l$, and $D^\alpha f(x) = \partial^{|\alpha|} f(x) / (\partial x_1^{\alpha_1} \times \cdots \times \partial x_k^{\alpha_k})$. (The partial derivative $D^\alpha f$ can be defined in the usual sense or in a weaker sense (see Stein [23, p. 180]).) Below we assume that $\sup_{\tau \in \mathscr{T}} \|m(\cdot, \tau)\|_{q,2,\chi} < \infty$ for some $q > (k+1)/2$.

Exponential Fourier functions can be used to obtain series expansions of the functions $\{m(\cdot, \tau)\}$. Let $(a, b)^k$ be an open bounded k-dimensional cube that contains the closure of $\mathscr{X}$. Following the approach of Edmunds and Moscatelli [11, p. 10], which uses Theorem 5 of Stein [23, p. 181], we extend the function $m(\cdot, \tau)$ on χ to a function $m^*(\cdot, \tau)$ on $(a, b)^k$ such that $m^*(\cdot, \tau)$ is periodic in each of its elements and

$$\|m(\cdot, \tau)\|_{\tilde{q},2,\chi} \leqslant \|m^*(\cdot, \tau)\|_{\tilde{q},2,(a,b)^k} \leqslant G \, \|m(\cdot, \tau)\|_{\tilde{q},2,\chi} \tag{3.2}$$

for all non-negative integers $\tilde{q} \leqslant q$, for some $G < \infty$ that does not depend on τ. The Fourier expansion of the periodic function $m^*(\cdot, \tau)$ gives the desired expansion of $m(\cdot, \tau)$ by restricting the domain of the expansion to $\mathscr{X}$. (Note that the extension operator used above is linear, so that Eq. (3.2) also holds with $m(\cdot, \tau)$ and $m^*(\cdot, \tau)$ replaced by $m(\cdot, \tau) - m(\cdot, \gamma)$ and $m^*(\cdot, \tau) - m^*(\cdot, \gamma)$, respectively, as used below.)

Call a k-vector of integers a *multi-index*. By Theorem 2 of Edmunds and Moscatelli [11, pp. 11, 25], a sequence of multi-indexes $\{\kappa(j) : j \geqslant 1\}$ can be constructed such that

$$\sup_{x \in (a,b)^k} \left| m^*(x, \tau) - \sum_{j=1}^{J} c_j(\tau) \, h_j(x) \right| \leqslant G^* J^{-q/k + 1/2 + \varepsilon} \, \|m^*(\cdot, \tau)\|_{q,2,(a,b)^k} \tag{3.3}$$

$\forall \varepsilon > 0$, for some $G^* < \infty$, where

$$h_j(x) = (b-a)^{-k/2} \, e^{2\pi i \kappa(j)'(x - a\mathbf{1})/(b-a)}$$

and

$$c_j(\tau) = \int_{(a,b)^k} m^*(x, \tau) \, \overline{h_j(x)} \, dx. \tag{3.4}$$

Here, $\mathbf{1}$ denotes a k-vector of ones and $\overline{h_j(x)}$ denotes the complex conjugate of $h_j(x)$. By (3.3) and the assumption $q > (k+1)/2$, we obtain the following pointwise convergent series expansion of $m(\cdot, \tau)$:

$$m(x, \tau) = \sum_{j=1}^{\infty} c_j(\tau)\, h_j(x) \qquad \forall x \in \mathscr{X}, \forall \tau \in \mathscr{T}. \tag{3.5}$$

Next we define the metric ρ on $\mathscr{T}$. For a function f on $\mathscr{Y} \subset R^k$, define

$$\|f\|_{p,\mathscr{Y}} = \left(\int_{\mathscr{Y}} |f(x)|^p\, dx\right)^{1/p} \qquad \text{for} \quad 1 \leqslant p < \infty. \tag{3.6}$$

Since the functions $\{h_j(\cdot): j \geqslant 1\}$ are orthonormal on $(a, b)^k$ (with respect to Lebesgue measure) and (3.2) holds with $\tilde{q} = 0$ and with $m(\cdot, \tau)$ and $m^*(\cdot, \tau)$ replaced by $m(\cdot, \tau) - m(\cdot, \gamma)$ and $m^*(\cdot, \tau) - m^*(\cdot, \gamma)$, respectively, we obtain

$$\begin{aligned}\left(\sum_{j=1}^{\infty} |c_j(\tau) - c_j(\gamma)|^2\right)^{1/2} &= \|m^*(\cdot, \tau) - m^*(\cdot, \gamma)\|_{2,(a,b)^k} \\ &\in [\|m(\cdot, \tau) - m(\cdot, \gamma)\|_{2,\mathscr{X}}, G\, \|m(\cdot, \tau) - m(\cdot, \gamma)\|_{2,\mathscr{X}}].\end{aligned} \tag{3.7}$$

Hence, we can define the metric ρ on $\mathscr{T}$ to be

$$\rho(\tau, \gamma) = \|m(\cdot, \tau) - m(\cdot, \gamma)\|_{2,\mathscr{X}} \tag{3.8}$$

and with this definition ρ is equivalent to (although not identical to) the metric ρ defined in (2.4).

Using Corollary 1, the following assumption is shown to imply that $\{v_n(\cdot)\}$ satisfies a CLT when indexed by the family of differentiable functions $\{m(\cdot, \tau)\}$ on $\mathscr{X}$:

Assumption C. (i) $\mathscr{X}$ is a bounded open subset of R^k with minimally smooth boundary.

(ii) $\sup_{\tau \in \mathscr{T}} \|m(\cdot, \tau)\|_{q,2,\mathscr{X}} < \infty$ for some $q > (k+1)/2$.

(iii) $\{X_{ni}\}$ is a NED triangular array of size -1 on $\{V_{ni}\}$, where $\{V_{ni}\}$ is some strong mixing double array of size -2.

(iv) $S(\tau, \gamma) = \lim_{n \to \infty} \operatorname{Cov}(v_n(\tau), v_n(\gamma))$ exists $\forall \tau, \gamma \in \mathscr{T}$.

THEOREM 4. *For $v_n(\cdot)$ and ρ as defined in* (2.1) *and* (3.8), *respectively, Assumption* C *implies that $v_n(\cdot)\, wv(\cdot)$, where $v(\cdot)$ is a mean zero Gaussian process with covariance function $S(\cdot, \cdot)$ whose sample paths lie in $\mathscr{U}_\rho(\mathscr{T})$ with probability one.*

Comments. 1. Since strong mixing triangular arrays are NED on themselves, Theorem 4 covers the case where $\{X_{ni}\}$ is strong mixing of size -2. In fact, in this case, the condition $q > (k+1)/2$ can be replaced by $q > k/2$. (This result is obtained by verifying Assumption A, rather than A1, in the proof of Theorem 4.)

2. The smoothness condition $q > (k+1)/2$ of Assumption C(ii) (or $q > k/2$ in the strong mixing case) can be compared with other conditions used in the literature to obtain CLTs for empirical processes indexed by smooth functions. In particular, if Assumptions C(i) and (ii) hold but with q unspecified, if $\{X_{ni}\}$ are i.i.d., and if one verifies Ossiander's [19, Theorem 3.1, p. 904) metric entropy with bracketing condition using Kolmogorov and Tihomirov's [16, Theorem XIII, p. 308] calculation of the (sup-norm) metric entropy numbers of $\{m(\cdot, \tau)\}$, then one needs the condition $q > k/2$. This is exactly the same condition as used here in the strong mixing $\{X_{ni}\}$ case, and it is only slightly weaker than the condition used in the NED case. (Note, however, that the result of Theorem 4 differs from Ossiander's result in that her result uses the $L^2(\mathbf{P})$ metric on $\mathscr{T}$, whereas Theorem 4 uses the $L^2(\mu)$ metric, where μ is Lebesgue measure on $\mathscr{X}$.)

3. For the case where $\mathscr{X}$ is a bounded interval (a, b) and $\{X_{ni}\}$ is strong mixing of size -2, the smoothness Assumption C(ii) can be relaxed. It can be replaced by:

C(ii)* $m(\cdot, \tau) \in \mathrm{Lip}(\alpha, D, (a, b))\ \forall \tau \in \mathscr{T}$ for some $\alpha \in (\frac{1}{2}, 1]$ and $D < \infty$.

(This result is proved by verifying Assumptions A and B using trigonometric series expansions and results of Zygmund [26, p. 135] and Dehling [8, Lemma 10.1, p. 428].)

The smoothness condition C(ii)* is exactly the same as the smoothness condition needed to verify Ossiander's [19, Theorem 3.1, p. 904) metric entropy with bracketing condition using Kolmogorov and Tihomirov's [16, Theorem XIII, p. 308] computation of the (sup norm) metric entropy numbers for Lipschitz functions on a bounded interval. Ossiander's bracketing condition yields an empirical process CLT for i.i.d. sequences of rv's. In addition, the condition $\alpha > \frac{1}{2}$ is needed by Dehling [8, Theorem 5, p. 399] to apply his CLT for Banach space-valued rv's to Lipschitz functions on a bounded interval. Dehling's CLT applies to sequences of weakly stationary strong mixing rv's of size -4.

4. Boundedness of X_{ni} can be a restrictive assumption in some applications of Theorem 4. It can be relaxed, however, in a way that is use-

ful in some applications. Suppose $m(X_{ni}, \tau)$ is of the form $m_1(X_{1ni}, \tau)\, X_{2ni}$, where $X_{ni} = (X'_{1ni}, X'_{2ni})'$ and $X_{1ni} \in R^{k_1}$. Let $\mathscr{X}_1$ be the space of possible values of X_{1ni}. Replace χ, $m(\cdot, \tau)$, and $k+1$ in Assumption C and (3.8) by $\mathscr{X}_1$, $m_1(\cdot, \tau)$, and k_1. Assume that the revised Assumption C holds, $\sup_{i \leqslant n, n \geqslant 1} E\, \|X_{2ni}\|^r < \infty$ for some $r > 2$, and $\{X_{ni}\}$ is a strong mixing triangular array of size $-2r/(r-2)$. Then, the conclusion of Theorem 4 still holds even though X_{ni} and $m(\cdot, \tau)$ may be unbounded.

To prove this, take the series expansion of $m(\cdot, \tau)$ to be given by X_{2ni} times a Fourier series expansion of $m_1(\cdot, \tau)$. Then, verify Assumption A rather than Assumption A1, since the series functions $h_j(\cdot)$ do not satisfy Assumption A1(ii) in this case. Next, note that the smoothness assumption B(iv) on $m(\cdot, \tau)$, which is violated in this case, can be eliminated and the result of Theorem 3 still holds, provided $\{X_{ni}\}$ satisfies the additional conditions introduced above.

5. If the support of the rv's $\{X_{ni}\}$ is not an open set, Theorem 4 still applies provided the functions $\{m(\cdot, \tau) \colon \tau \in \mathscr{T}\}$ are defined on an open bounded set $\mathscr{X}$ that contains the support of $\{X_{ni}\}$. On the other hand, if the functions $\{m(\cdot, \tau) \colon \tau \in \mathscr{T}\}$ are only defined on the support of the $\{X_{ni}\}$ and certain elements of X_{ni} only take on a finite number of values, then $\mathscr{X}$ cannot be an open set. The latter problem can be circumvented in this case by writing $v_n(\cdot)$ as the sum of several empirical processes based on X_{ni} vectors of lower dimension with the discrete elements of X_{ni} eliminated.

Proof of Theorem 4. We show that $C \Rightarrow A1$ and $C \Rightarrow B$. Corollary 1 then gives the result.

Under Assumptions C(i) and (ii), Assumption A1(i) holds by the argument given in Eqs. (3.3)–(3.5) above. Next we show Assumption A1(ii) holds. Since $(a, b)^k$ is open and convex, the mean value theorem gives: $\forall x, y \in (a, b)^k$,

$$\begin{aligned} |h_j(x) - h_j(y)| &\leqslant \sup_{x^* \in (a,b)^k} \left\| \frac{\partial}{\partial x} h_j(x^*) \right\| \|x - y\| \\ &\leqslant 2\pi (b-a)^{-1-k/2} \sum_{l=1}^{k} |\kappa_l(j)|\, \|x - y\|. \end{aligned} \tag{3.9}$$

Since the multi-indexes $\{\kappa(j)\}$ considered by Edmunds and Moscatelli [11, p. 11] and used in (3.3)–(3.5) above satisfy $|\kappa_l(j)| \leqslant D^{j^{1/k}}\ \forall l = 1, ..., k$ for some $D < \infty$, $h_j(\cdot)$ is in Lip$(1, B_j, (a, b)^k)$ with $B_j = 2\pi(b-a)^{-1-k/2}\, kDj^{1/k}$. In addition, $|h_j(x)|$ is bounded by $(b-a)^{-k/2}$ for all $x \in \mathscr{X}$ and $j \geqslant 1$. Thus, Assumption A1(ii) holds for all $r < \infty$.

To establish A1(iii), let $a_j = j^{-2q/k + \varepsilon}$ for some $\varepsilon \in (0, -1 + (2q-1)/k)$.

Then,

$$\sup_{\tau \in \mathscr{T}} \sum_{j=J}^{\infty} |c_j(\tau)|^2/a_j = \sup_{\tau \in \mathscr{T}} \sum_{j=J}^{\infty} j^{2q/k-\varepsilon} |c_j(\tau)|^2$$

$$\leqslant J^{-\varepsilon} \sup_{\tau \in \mathscr{T}} \sum_{j=1}^{\infty} j^{2q/k} |c_j(\tau)|^2 = o(1) \quad \text{as} \quad J \to \infty, \tag{3.10}$$

since $\sup_{\tau \in \mathscr{T}} \sum_{j=1}^{\infty} j^{2q/k} |c_j(\tau)|^2 < \infty$ by (3.11) below. For $q > (k+1)/2$ and ε as defined above, $\sum_{j=1}^{\infty} a_j B_j \propto \sum_{j=1}^{\infty} j^{-(2q-1)/k+\varepsilon} < \infty$, and hence, Assumption A1(iii) holds.

To show $\sup_{\tau \in \mathscr{T}} \sum_{j=1}^{\infty} j^{2q/k} |c_j(\tau)|^2 < \infty$, we use Assumption C(ii) and (3.2). Let (f, g) denote $\int_{(a,b)^k} f(x) \overline{g(x)}\, dx$. We have

$$\infty > G \sup_{\tau \in \mathscr{T}} \|m(\cdot, \tau)\|_{q,2,\mathscr{X}}$$

$$\geqslant \sup_{\tau \in \mathscr{T}} \|m^*(\cdot, \tau)\|_{q,2,(a,b)^k}$$

$$= \sup_{\tau \in \mathscr{T}} \left(\sum_{|\alpha| \leqslant q} \sum_{j=1}^{\infty} |(D^\alpha m^*(\cdot, \tau), h_j(\cdot))|^2 \right)^{1/2}$$

$$= \sup_{\tau \in \mathscr{T}} \left(\sum_{|\alpha| \leqslant q} \sum_{j=1}^{\infty} \left(\frac{2\pi}{b-a}\right)^{2|\alpha|} \prod_{v=1}^{k} \kappa_v(j)^{2\alpha_v} |(m^*(\cdot, \tau), h_j(\cdot))|^2 \right)^{1/2}$$

$$\geqslant \min\left(\left(\frac{2\pi}{b-a}\right)^q, 1\right) \left(\sup_{\tau \in \mathscr{T}} \sum_{j=1}^{\infty} \max_{1 \leqslant v \leqslant k} \kappa_v(j)^{2q} |c_j(\tau)|^2 \right)^{1/2}$$

$$\geqslant B^* \min\left(\left(\frac{2\pi}{b-a}\right)^q, 1\right) \left(\sup_{\tau \in \mathscr{T}} \sum_{j=1}^{\infty} j^{2q/k} |c_j(\tau)|^2 \right)^{1/2} \tag{3.11}$$

for some constant $B^* < \infty$. The last inequality holds because $|\kappa(j)| = \sum_{v=1}^{k} |\kappa_v(j)| \approx j^{1/k}$ (see Edmunds and Moscatelli [11, p. 11]) implies that $\max_{1 \leqslant v \leqslant k} |\kappa_v(j)| \approx j^{1/k}$ (where $f(j) \approx g(j)$ means that $0 < \underline{\lim}_{j \to \infty} |f(j)/g(j)| \leqslant \overline{\lim}_{j \to \infty} |f(j)/g(j)| < \infty$).

Assumption A1(iv) holds for some r sufficiently large by C(iii).

Next we show $C \Rightarrow B$. Assumptions B(i) and C(iv) are equivalent. Assumption B(ii) holds for all $r < \infty$ since $\mathscr{X}$ is bounded. Assumption B(iii) holds for some r sufficiently large by Assumption C(iii). To show Assumption B(iv), it suffices to show $m^*(\cdot, \tau) \in \text{Lip}(1, C, (a, b)^k)\ \forall \tau \in \mathscr{T}$ for some $C < \infty$. This follows by the same argument as in (3.9) using the uniform bound on the first derivatives of $m^*(\cdot, \tau)$ given by Assumption C(ii) and (3.2). ∎

4. Example

This section briefly illustrates the use of the results given above in establishing the asymptotic normality and efficiency of a semiparametric estimator. We consider a weighted least squares (LS) estimator of a dynamic nonlinear regression model with weights that adapt to the form of heteroskedasticity present. Carroll [7] and Robinson [22] have considered this estimator in non-dynamic linear regression models. The model is

$$Y_i = f(Y_{i-1}, ..., Y_{i-p}, Z_i, \theta_0) + U_i \qquad \text{for} \quad i = 1, ..., n, \tag{4.1}$$

where Y_i, $U_i \in R$, $Z_i \in R^{\xi}$, and $\theta_0 \in \Theta \subset R^{\gamma}$. The errors $\{U_i\}$ are independent with the conditional mean of U_i given $(Y_{i-1}, ..., Y_{i-p}, Z_i)$ equal to 0 and the conditional variance function $\tau_0(\cdot)$ of U_i defined by

$$\tau_0(Z_i) = E(U_i^2 \mid Z_i) = E(U_i^2 \mid Y_{i-1}, ..., Y_{i-p}, Z_i). \tag{4.2}$$

The regressors $\{Z_i\}$ may be fixed or random. θ_0 is an unknown parameter.

The parameter θ_0 is estimated using a preliminary estimator $\hat{\tau}(\cdot)$ of $\tau_0(\cdot)$:

$$\hat{\theta} \text{ minimizes } \sum_{i=1}^{n} (Y_i - f_i(\theta))^2/\hat{\tau}(Z_i) \tag{4.3}$$

over $\theta \in \Theta$, where $f_i(\theta)$ denotes $f(Y_{i-1}, ..., Y_{i-p}, Z_i, \theta)$. Under suitable conditions (e.g., see Andrews [2]), this estimator of θ_0 is consistent and satisfies

$$\sqrt{n}\,(\hat{\theta} - \theta_0) = \left(\frac{1}{n}\sum_{i=1}^{n} E\frac{\partial}{\partial\theta} f_i(\theta_0)\frac{\partial}{\partial\theta'} f_i(\theta_0)/\tau_0(Z_i)\right)^{-1} \times \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\hat{\tau}(Z_i) + o_p(1), \tag{4.4}$$

where the inverted matrix is $O(1)$ as $n \to \infty$. The linear approximation (4.4) and the CLT give the asymptotic normality of $\sqrt{n}\,(\hat{\theta} - \theta_0)$ if it can be shown that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\hat{\tau}(Z_i) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\tau_0(Z_i) \xrightarrow{p} 0. \tag{4.5}$$

If the errors are normally distributed, then (4.4) and (4.5) also combine to establish the asymptotic efficiency of $\hat{\theta}$, since $\hat{\theta}$ has the same asymptotic distribution as the weighted LS "estimator" that uses $\tau_0(\cdot)$ to form the weights.

Define an empirical process $v_n(\cdot)$ as

$$v_n(\tau) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} U_i \frac{\partial}{\partial\theta} f_i(\theta_0)/\tau(Z_i) \qquad \text{for} \quad \tau \in \mathscr{T}, \tag{4.6}$$

where $\mathscr{T}$ is a class of smooth functions $\tau(\cdot)$ for which $E\,\|U_i(\partial/\partial\theta) f_i(\theta_0)/\tau(Z_i)\|^2 < \infty$. Note that $EU_i(\partial/\partial\theta) f_i(\theta_0)/\tau(Z_i) = 0$ $\forall\tau \in \mathscr{T}$, since U_i has conditional mean zero.

The results of this paper (in particular, Theorem 4 or Comment 5 following Theorem 4) can be used to show that $\{v_n(\cdot): n \geqslant 1\}$ is stochastically equicontinuous. Equation (4.5) then follows provided $\hat{\tau}$ converges in probability to τ_0 with respect to the appropriate metric. For example, if Theorem 4 or Comment 5 following Theorem 4 is used, then $\hat{\tau} \xrightarrow{p} \tau_0$ if

$$\int_{\mathscr{Z}} (\hat{\tau}(z) - \tau_0(z))^2 \, dz \xrightarrow{p} 0 \qquad \text{as} \quad n \to \infty \tag{4.7}$$

and $\hat{\tau}(\cdot)$ and $\tau_0(\cdot)$ are bounded away from zero, where $\mathscr{Z}$ denotes the (bounded) set of possible realizations of $\{Z_i: i \geqslant 1\}$.

In sum, the stochastic equicontinuity of an empirical process can be used very effectively in establishing the asymptotic normality and efficiency of semiparametric estimators. See Andrews [2, 3] for the application of empirical process results, such as those of the present paper, to a broad class of semiparametric estimators and to tests based on them.

References

1. Andrews, D. W. K. (1985). A nearly independent, but non-strong mixing, triangular array. *J. Appl. Probab.* **22** 729–731.
2. Andrews, D. W. K. (1989). Asymptotics for Semiparametric Econometric Models. I. Estimation and Testing. Cowles Foundation Discussion Paper No. 908R, Yale University.
3. Andrews, D. W. K. (1989). Asymptotics for Semiparametric Econometric Models. II. Stochastic Equicontinuity and Nonparametric Kernel Estimation. Cowles Foundation Discussion Paper No. 909R, Cowles Foundation, Yale University.
4. Andrews, D. W. K., and Pollard, D. (1990). A functional central limit theorem for strong mixing stochastic processes. Cowles Foundation Discussion Paper No. 951, Cowles Foundation, Yale University.
5. Bierens, H. J. (1981). *Robust Methods and Asymptotic Theory in Nonlinear Econometrics.* Lecture Notes in Economics and Mathematical Systems, Vol. 192. Springer-Verlag, New York.
6. Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.
7. Carroll, R. J. (1982). Adapting for heteroscedasticity in linear models. *Ann. Statist.* **10** 1224–1233.
8. Dehling, H. (1983). Limit theorems for sums of weakly dependent Banach space valued random variables. *Z. Wahrsch. Verw. Gebiete* **63** 393–432.

9. DEHLING, H., AND PHILIPP, W. (1982). Almost sure invariance principles for weakly dependent random vectors. *Ann. Probab.* **10** 689–701.
10. DUDLEY, R. M. (1985). An extended Wichura theorem, definitions of Donsker class, and weighted empirical distributions. In *Lecture Notes in Math.*, Vol. 1153, pp. 141–178. Springer-Verlag, New York.
11. EDMUNDS, D. E., AND MOSCATELLI, V. B. (1977). Fourier approximations and embeddings in Sobolev space. *Dissertationes Math.* **145** 1–46.
12. GALLANT, A. R. (1987). *Nonlinear Statistical Models.* Wiley, New York.
13. GALLANT, A. R., AND WHITE, H. (1988). *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models.* Basil Blackwell, New York.
14. HOFFMANN-JORGENSEN, J. (1985). Stochastic processes on Polish spaces. Unpublished manuscript, University of Aarhus, Denmark.
15. IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Theory Probab. Appl.* **7** 349–382.
16. KOLMOGOROV, A. N., AND TIHOMIROV, V. M. (1961). ε-entropy and ε-capacity of sets in functional spaces. *Amer. Math. Soc. Transl. Ser. 2* **17** 277–364.
17. MCLEISH, D. L. (1975). A maximal inequality and dependent strong laws. *Ann. Probab.* **3** 829–839.
18. MCLEISH, D. L. (1975). Invariance principles for dependent variables. *Z. Wahrsch. Verw. Gebiete* **32** 165–178.
19. OSSIANDER, M. (1987). A central limit theorem under metric entropy with L_2 bracketing. *Ann. Probab.* **15** 897–919.
20. POLLARD, D. (1989). A maximal inequality for sums of independent processes under a bracketing condition. Unpublished manuscript, Department of Statistics, Yale University.
21. POLLARD, D. (1990). *Empirical Processes: Theory and Applicátions.* CBMS Regional Conference Series in Probability and Statistics, Vol. 2, Institute of Mathematical Statistics, Hayward, CA.
22. ROBINSON, P. M. (1987). Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form. *Econometrica* **55** 875–891.
23. STEIN, E. M. (1970). *Singular Integrals and Differentiability Properties of Functions.* Princeton Univ. Press, Princeton, NJ.
24. WITHERS, C. S. (1981). Central limit theorems for independent variables I. *Z. Wahrsch. Verw. Gebiete* **57** 509–534.
25. WOOLDRIDGE, J. (1986). *Asymptotic Properties of Econometric Estimators.* Ph.D. dissertation, Department of Economics, University of California, San Diego.
26. ZYGMUND, A. (1955). *Trigonometrical Series.* Dover, New York. First published, 1935.